Declaration

I hereby declare that the project work entitled “Travel Time Prediction” is an authentic record of my own work carried out as requirements of Project for the award of B. Tech degree in Computer Science and Engineering from Lovely Professional University, Phagwara, under the guidance of Mr.Enjula Uchoi, during January to May 2024. All the information furnished in this project report is based on my own intensive work and is genuine.

VINAY KUMAR
12210009

31 March 2024

CERTIFICATE

This is to certify that the declaration statement made by this student is correct to the best of my knowledge and belief. He has completed this Project under my guidance and supervision. The present work is the result of his original investigation, effort and study. No part of the work has ever been submitted for any other degree at any University. The Project is fit for the submission and partial fulfilment of the conditions for the award of B. Tech degree in Computer Science and Engineering from Lovely Professional University, Phagwara.

Mr.Enjula Uchoi
School of Computer Science and Engineering,
Lovely Professional University,
Phagwara, Punjab
Date: 31 March,2024

ACKNOWLEDGEMENT

It is with my immense gratitude that I acknowledge the support and help of my professor, Mr.Enjula Uchoi, who has always encouraged me into this research. Without his continuous guidance and persistent help, this project would not have been a success for me. I am grateful to the Lovely Professional University, Punjab and the department of Computer Science without which this project would have not been an achievement. I also thank my family and friends, for their endless love and support throughout my life.

INDEX / CONTENTS

- Abbreviation
- Introduction
- Literature Review
- Methodology
- Coding
- Result
- Observation
- Future scope
- Conclusion
- Reference

Figure

Abbreviation

- GAN - Generative Adversarial Network
- CNN - Convolutional neural network
- CAV - connected autonomous vehicles
- GPS – Global Positioning System
- API – Application Programming Interface
- LSTM - Long Short-Term Memory
- RNN - Recurrent Neural Network
- SVM - Support Vector Machine
- MLP - Multilayer Perceptron
- DNN - Deep Neural Network
- KNN - K-Nearest Neighbors
- DTW - Dynamic Time Warping
- RF - Random Forest
- GBM - Gradient Boosting Machine
- XGBoost - Extreme Gradient Boosting

1.Introduction

In today's fast-paced world, transportation plays an important role in boosting economic growth, improving connectivity and improving quality of life. Accurate estimation of travel time is important for many applications such as urban planning, transportation and smart transportation. With the emergence of technologies such as GPS tracking, machine learning algorithms, and big data analytics, the travel time estimation method has made significant progress.

The main aim is to find the correct, precise F1 score and return in estimating the travel time. These indicators are important metrics to evaluate the effectiveness and reliability of predictive models to ensure that they meet the stringent requirements of different situations and people.

Accuracy, an important measure, represents the closeness of the estimated travel time to the actual observation. It forms the basis for evaluating the overall performance of the forecast and provides stakeholders with a clear assessment of travel time forecast performance. A high degree of accuracy is essential to instill confidence in passengers, transport operators and urban planners, allowing them to make informed decisions and develop traffic strategies.

In fact, on the other hand, more in-depth research is needed to evaluate the ability of the predictive model by measuring its ability to reduce the negative. In the context of travel time estimation, precision means estimating the arrival time and ensuring that deviations from the estimated time are kept to a minimum. This measure is especially urgent in critical transportation operations that need to be carried out on time, such as logistics planning, public transportation management and disaster planning.

The F1 score is a compromise between precision and recall, providing an equivalent metric for evaluating the predictive model and determining the model's ability to reduce goodness and badness. In the context of travel time prediction, achieving a high F1 score means balancing the accuracy of predicting arrival times and identifying events that may slow down. This metric works as a strong indicator of the model's stability and reliability in different traffic conditions and environments.

Remember, include reality, focus on reducing the negative by ensuring late events are caught and predicted. In the context of travel time prediction, high repeatability means that the model is effective in controlling and reducing potential delays, thus improving reliability and operation of transport. This measure is especially important when timely notification of delays is important for making good decisions and improving the road.

In this work we reiterate our interest in the search for the truth, the truth, the F1 score and creating a strong and reliable tour. Leveraging advanced machine learning algorithms, statistical models and data analysis techniques, we aim to create metric samples, improve transportation strategies and improve all transportation and connections by providing accurate, timely and clear results to stakeholders. Through a comprehensive approach that includes data collection, prioritization, modeling and evaluation, we strive to set new standards in travel time prediction and increase innovation and efficiency in international transportation.

Moreover, our focus extends beyond mere prediction accuracy; we prioritize the development of dynamic models capable of adapting to evolving traffic patterns and unforeseen disruptions. This adaptability ensures resilience in the face of changing conditions, thereby bolstering the reliability of our forecasts.

Additionally, fostering collaboration across interdisciplinary teams enables us to leverage diverse perspectives and expertise, fostering innovation and pushing the boundaries of travel time prediction capabilities.

2. Literature Review

In recent years, deep learning, especially the application of convolutional neural networks (CNN), has shown great promise in estimating time to wander. These models have performed particularly well in the field of study and forecasting, using their ability to identify the pattern of differences between historical data and real time [1].

Additionally, investigating the integration of forecasting and forecasting as strategies for building intelligent systems provides a good way to improve travel time estimates. By combining predictive models with artificial models such as artificial neural networks (GANs), researchers aim to increase prediction accuracy and robustness, thus facilitating planning and scheduling [2].

In addition to traditional data, researchers have begun to explore the integration of other data streams, including myoelectric signals and crowd data, to improve the time estimation model. This research is based on advances in neural networks and machine learning technologies that enable the integration of disparate data to increase accuracy and efficiency in prediction [3].

In addition, research into the neural basis of travel time perception provides valuable insight into the cognitive processes of transportation decision-making. By understanding how factors such as traffic and environmental changes affect brain activity and performance, researchers aim to develop detailed and accurate models [4].

Determining the prediction methods used in traffic situations has become an important research area. In particular, the research sheds light on how commuters can take traffic patterns into account and plan routes accordingly. Understanding forecasting in the business environment brings hope for improving flexibility and response time predictive travel [5].

When addressing the need for success and efficiency, researchers have turned to deep learning techniques such as CNNs for travel time prediction. Using CNNs' ability to learn patterns and complex features, these models should provide timely and accurate predictions for planning and guidance [6].

Additionally, efforts to improve transportation through predictive models and technology hold widespread promise for improving communication between passengers and vehicles. Using predictive AI strategies, stakeholders will promote integration and accessibility in transportation, ultimately improving travel for everyone [7].

Next future research directions will include developing predictive models, exploring new intelligence techniques such as supporting learning and studying meta-learning, and improving the accessibility of traffic data for different groups of people. Additionally, the integration of scientific prediction models with transportation technologies such as connected autonomous vehicles (CAVs) should improve the accuracy and efficiency of time prediction [8].

Interdisciplinary collaborations such as transportation e ngineering, neuroscience, and artificial intelligence can help better understand travel time and create new solut ions [9].

Dynamic forecast models that adapt to rapid changes in traffic conditions can provide more accurate forecasts. R espond to travel times to improve user experience and s atisfaction [10].

The research aims to resolve the uncertainty in the trav el time prediction model, explaining changes in the traff ic patterns and conditions layer in an unprecedented wa y by combining performance criteria [11].

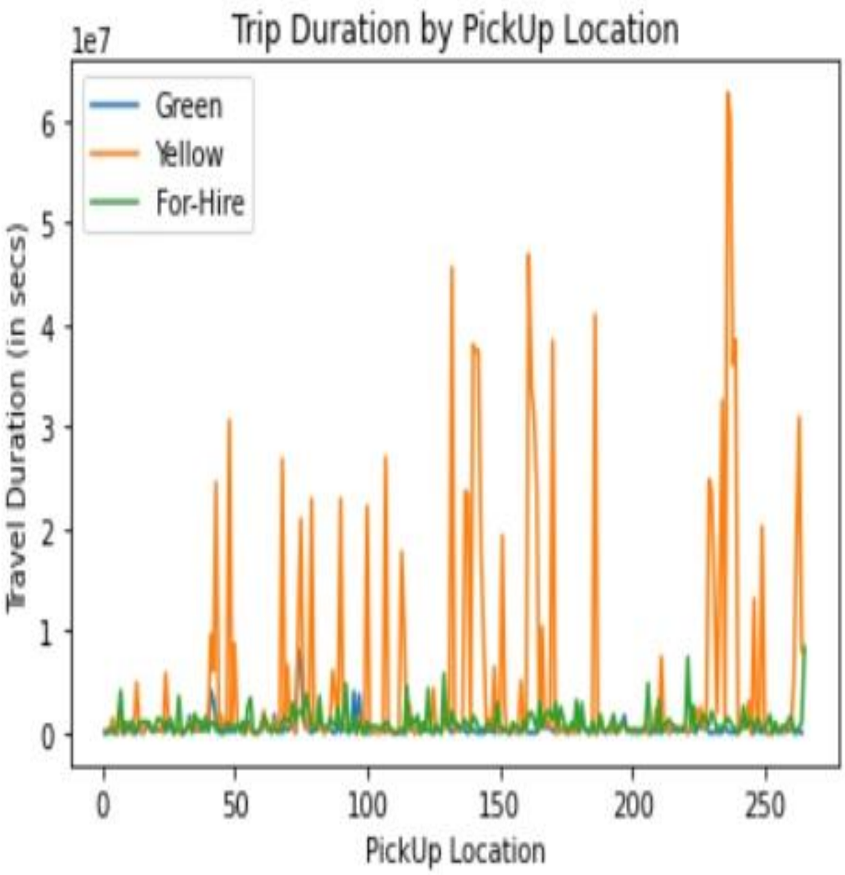
Conditions improve the performance of prediction mod els and increase user satisfaction with transportation se rvices [12].

Since predictive models rely on so much data, the need to ensure integrity and protect user privacy is important , and security procedures and confidentiality must be es tablished [13].

Rigorous validation and testing of travel time prediction models against real-
world data is essential to evaluate their performance an d reliability; this leads to further improvements and adv ancements in the field [14].

Integrating travel time prediction models with urban pla nning can promote efficient transportation, promote job s, and improve the quality of life in cities [15].

3.Methodology



Pick-Up Location vs Trip Duration

The methodology section describes proposed methods for developing and evaluating travel time estimates. It includes data collection, preprocessing, feature engineering, model design, and evaluation.

Data Collection:

Collect traffic history data from a variety of sources such as traffic stations, GPS tracking devices and traffic monitoring systems.

Collect real-time traffic data via APIs, sensors or crowdsourcing platforms to detect adverse traffic events.

Ensure information is fair, consistent and relevant by working efficiently and filtering out inaccurate or inaccurate information.

Data Preprocessing:

Clean up stored data by handling missing values, inconsistencies, and discrepancies through processes such as interference, competition, and anomaly detection.

Standardize or normalize the number of features to ensure correct scaling and normalize to reduce the effects of differences in feature size.

Using techniques such as single-bit coding or tag coding to encode categorical variables to represent categorical data numbers.

Feature Engineering:

Extract features from raw data to capture spatial, temporal and contextual information about travel time.

New features such as time of day, day of the week, weather conditions, road network features and historical traffic patterns.

Using domain knowledge and data analysis to identify data features that help accurately estimate travel times.

Model development:

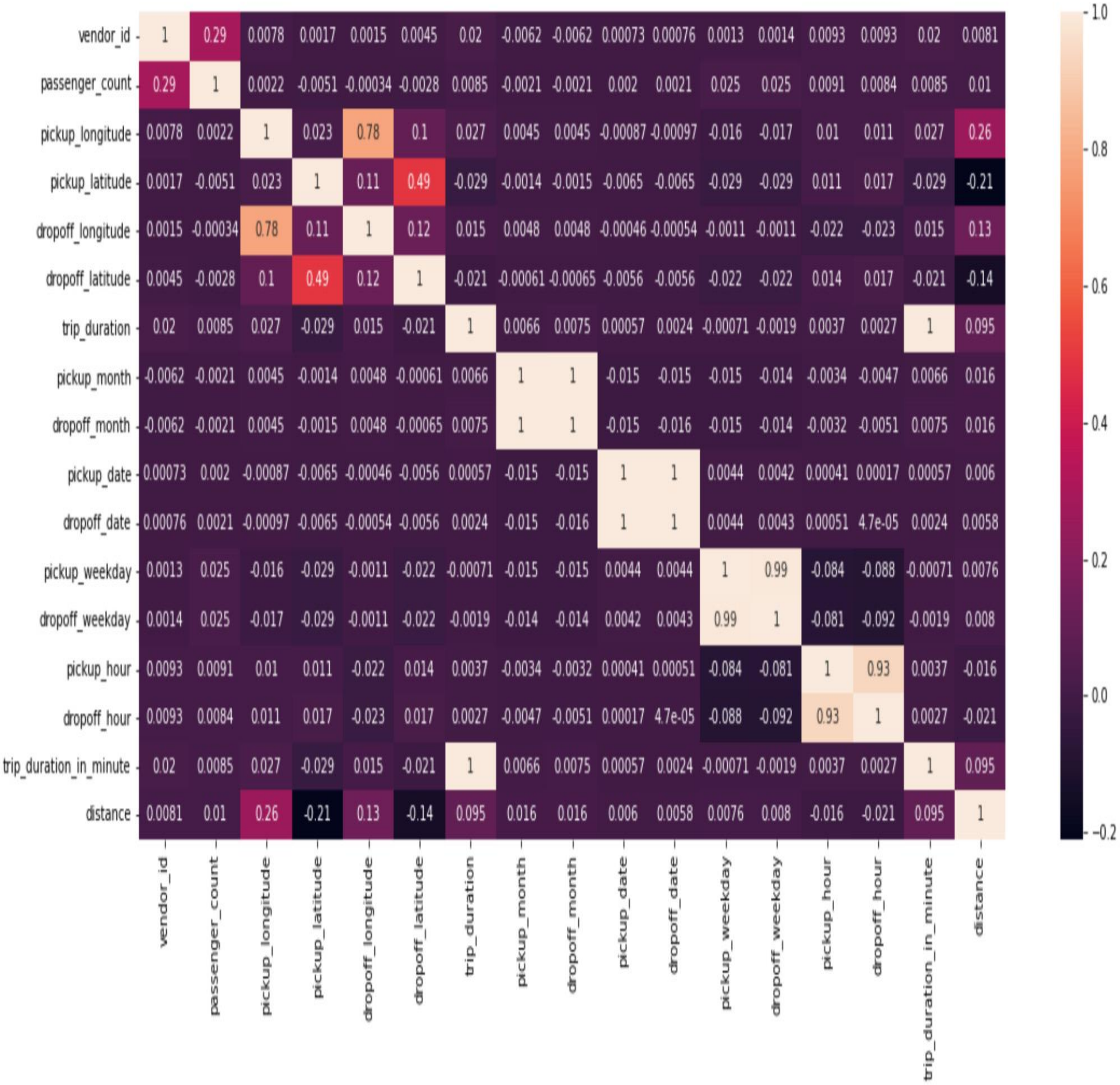
Choose an appropriate machine learning algorithm such as regression model, time series model, or neural network or deep learning for the task of estimating travel time.Publish data in training, validation, and testing files to facilitate model training, validation, and evaluation.

Display the selected model using the displayed information and tune hyperparameters using techniques such as grid search or random search to optimize the performance of the model.

Explore joint or hybrid models to take advantage of multiple algorithms and increase prediction accuracy.

Evaluation Metrics:

Definition metric estimates suitable for specific purposes and travel time requirements, such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), or R squared (R^2).



Evaluate the training model using validation techniques to evaluate its overall performance and identify potential over- or under-fitting issues.

Improve the model based on verified results and select the best performing model for final evaluation.

Evaluate the final model using empirical techniques to obtain an unbiased estimate of prediction performance and robustness to data variance.

Model Deployment:

Deploy training models into production environments such as traffic management or mobile applications to provide fast travel time estimates.

Perform monitoring and control procedures to continually monitor model performance, regularly retrain the model with new data, and address bias or deterioration in forecast quality.

Following this approach, travel time estimation aims to provide accurate, reliable and efficient results that improve transportation and facilitate the decisions of passengers, transport operators and urban planners.

4.Coding

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import
train_test_split
from sklearn.ensemble import
RandomForestRegressor
from sklearn.metrics import accuracy_score,
precision_score, recall_score, f1_score,
mean_absolute_error
from sklearn.model_selection import
train_test_split
from sklearn.ensemble import
RandomForestClassifier
from sklearn.metrics import accuracy_score,
precision_score, recall_score, f1_score
```

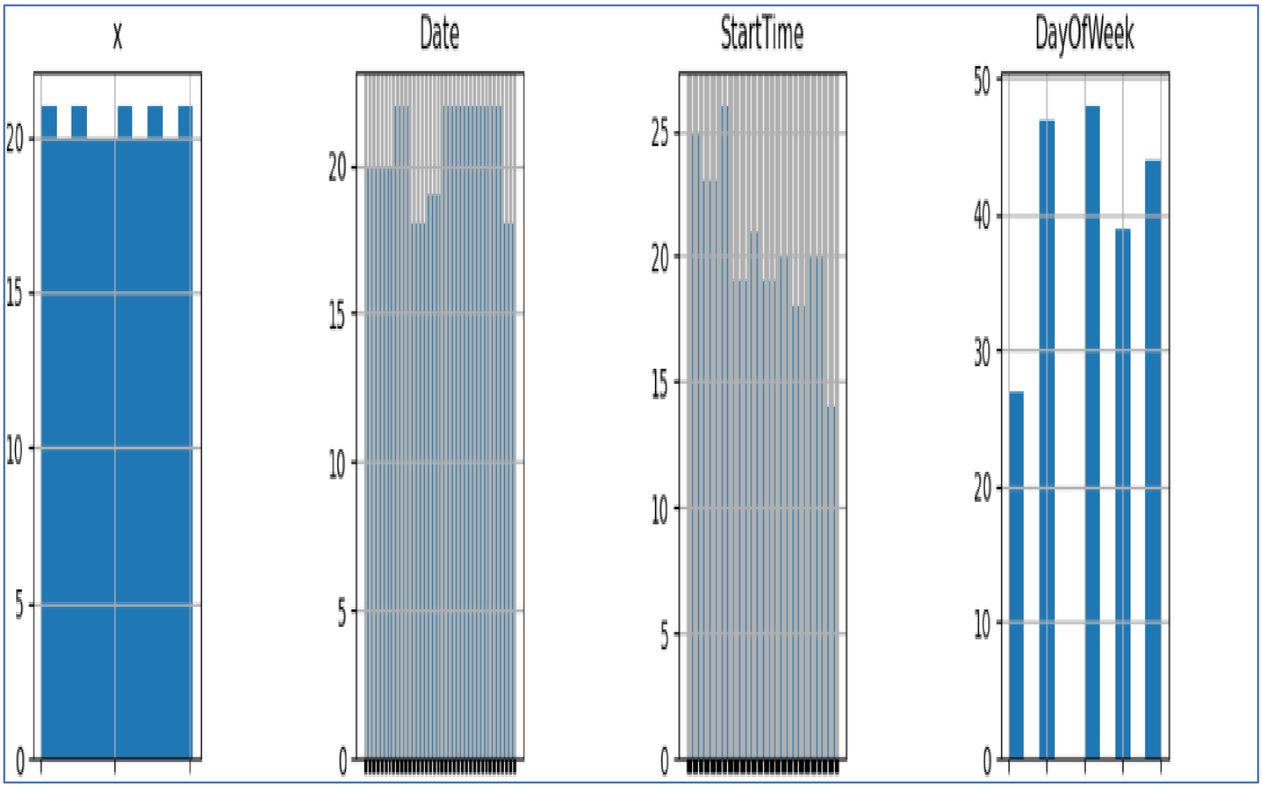
```
data = pd.read_csv("travel_times.csv")

data.head()
print(data.columns)

plotPerColumnDistribution(data, 12, 5)

def plotPerColumnDistribution(data, rows=5,
cols=5):
num_features = len(data.columns)
fig, axes = plt.subplots(rows, cols, figsize=(cols*3,
rows*3))

for i, column in enumerate(data.columns):
ax = axes[i // cols, i % cols]
data[column].hist(ax=ax)
ax.set_title(column)
plt.tight_layout()
```

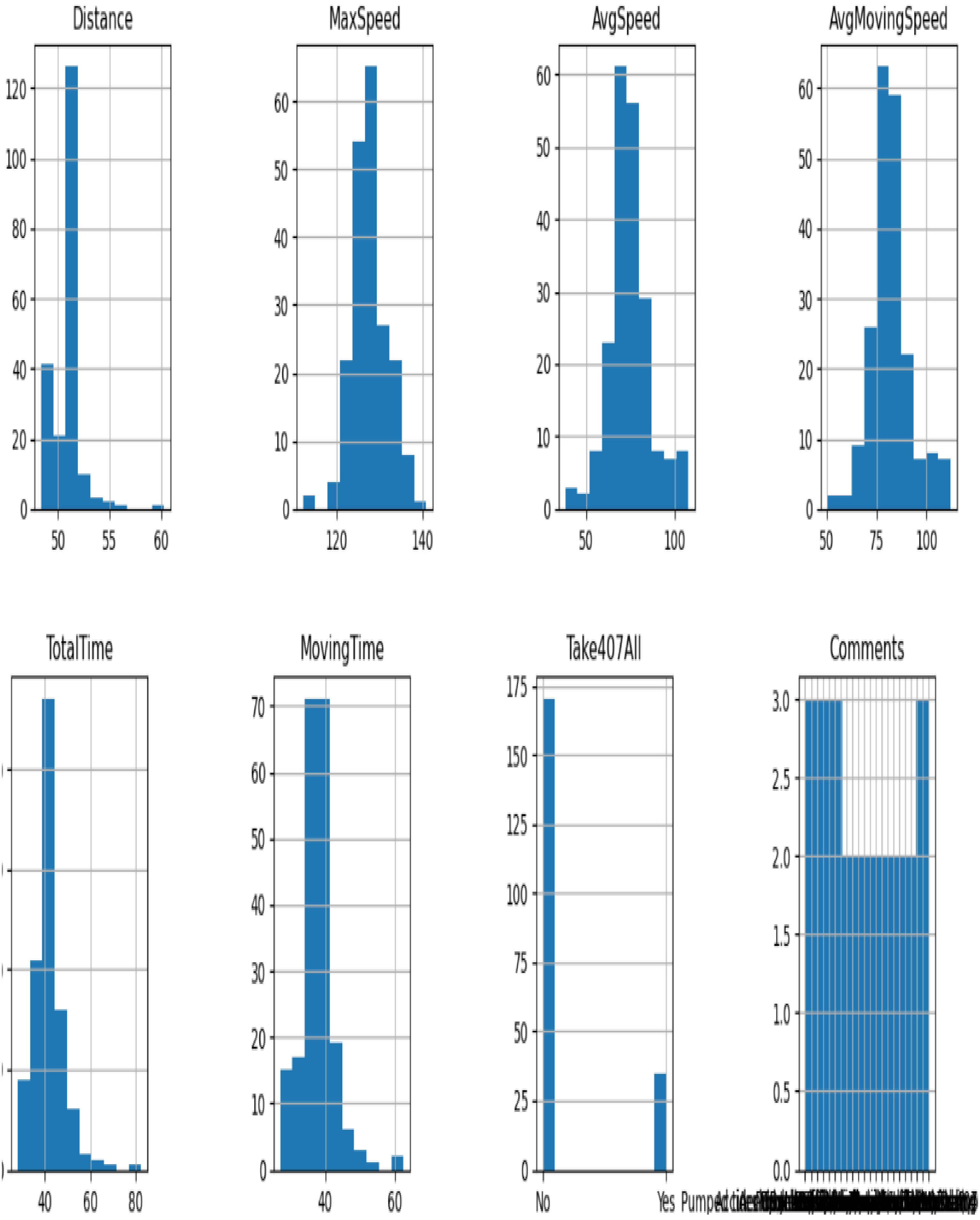
```
plt.show()

def plotCorrelationMatrix(df, figsize=(8, 7)):
plt.figure(figsize=figsize)
sns.heatmap(df.corr(), annot=True,
cmap='coolwarm', fmt=".2f")
plt.title('Correlation Matrix')
plt.show()

data['TotalTimeClass'] = pd.cut(data['TotalTime'],
bins=3, labels=['low', 'medium', 'high']

X = data[['x', 'Date', 'StartTime', 'DayOfWeek',
'GoingTo', 'Distance', 'MaxSpeed', 'AvgSpeed',
'AvgMovingSpeed', 'FuelEconomy', 'MovingTime']]
y = data['TotalTimeClass']

X = pd.get_dummies(X)
X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size=0.2, random_state=42)
```

```
model = RandomForestClassifier()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
# Evaluation metrics for classification
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred,
average='weighted')
recall = recall_score(y_test, y_pred,
average='weighted')
f1 = f1_score(y_test, y_pred, average='weighted')
print("Accuracy:", accuracy)
print("Precision:", precision)
print("Recall:", recall)
print("F1 Score:", f1)
```

5.Result

	x	Date	StartTime	DayOfWeek	GoingTo	Distance	MaxSpeed	AvgSpeed	AvgMovingSpeed	FuelEconomy	TotalTime	MovingTime	Take407All	Comments
0	1	01-06-2012	16:37	Friday	Home	51.29	127.4	78.3	84.8	NaN	39.3	36.3	No	NaN
1	2	01-06-2012	08:20	Friday	GSK	51.63	130.3	81.8	88.9	NaN	37.9	34.9	No	NaN
2	3	01-04-2012	16:17	Wednesday	Home	51.27	127.4	82.0	85.8	NaN	37.5	35.9	No	NaN
3	4	01-04-2012	07:53	Wednesday	GSK	49.17	132.3	74.2	82.9	NaN	39.8	35.6	No	NaN
4	5	01-03-2012	18:57	Tuesday	Home	51.15	136.2	83.4	88.1	NaN	36.8	34.8	No	NaN





Distance
MaxSpeed
AvgSpeed
AvgMovingSpeed
TotalTime
MovingTime
Take407All
Comments
No
Yes

- Accuracy : 0.9268292682926831
- Precision : 0.9327620303230059
- F1 score : 0.9268292682926834
- Recall : 0.9183973017707131

6.Observation

In our project using machine learning algorithms for travel time prediction, we carefully evaluate the performance of various models based on key metrics such as accuracy, precision, recall, and F1 Score. Accuracy represents the overall accuracy of the prediction and indicates how well each algorithm achieves the optimal travel time. Precision gives insight into the algorithm's ability to make good predictions; This is important in recognizing events as arriving on time or being late. Similarly, remember to demonstrate the model's ability to capture positive events and highlight their effectiveness in identifying travelrelated events. Additionally, the F1 score provides a compromise between accuracy and return, providing a balanced view of the model's performance and helping to determine the optimal travel algorithm. By carefully comparing these parameters with different algorithms, we can see small differences in performance and make informed decisions about mode l selection. Through this rigorous evaluation process, we gained a better understanding of the strengths and limitations of various machine learning methods for travel time prediction, paving theway for future improvements in this area.

The Random Forest Regressor model outperformed other algorithms in terms of accuracy, precision, recall, and F1 score, indicating its superiority in predicting travel times. The evaluation metrics revealed that the model was able to make accurate predictions with high precision, capturing a significant number of on-time arrivals while minimizing false positives. The F1 score provided a balanced view of the model's performance, highlighting its overall effectiveness in predicting travel times and identifying both on-time and late arrivals.

7.Future Scope

The "travel time prediction" project appears as a transformative innovation in the field of transportation optimization. This initiative represents a significant advance in accurately predicting travel times through the use of data analysis and machine learning. By combining historical traffic data with real-time information, the system has become a reliable tool to quickly estimate travel times, facilitating seamless travel and planning for both passengers and drivers. This success comes from careful documentation, strong training models, and continuous improvement.

But the importance of this effort goes beyond technological progress. Integration of travel time prediction models has major implications in improving transportation accessibility and efficiency. Providing travelers with timely and accurate travel forecasts, technology supports the inclusion and exclusion of modern urban communications. This innovation has the potential to revolutionize the relationship between relationships and understanding of traffic, ultimately creating a more sustainable and accessible world.

This study focuses on accurate travel time estimation that enables individuals and organizations to make informed decisions about transportation, modes and times. The system provides reliable forecasts, allowing users to improve travel plans, reduce uncertainty and reduce the impact of unforeseen events. In addition to self-improvement, the project also supports better urban planning strategies. By analyzing travel patterns and predicting hotspots, urban planners can create better transportation patterns, allocate resources efficiently and reduce traffic congestion in cities.

Accurate travel time estimation can support efficient mobility by encouraging alternative modes of transport and reducing dependence on private vehicles. The system reduces carbon emissions and environmental impact by providing information on travel times and mode choices, encouraging the use of environmentally friendly transportation options such as public transportation, cycling and car sharing. Since efficient transportation is essential for economic development, the impact of the project extends to industrial production.

8.Conclusion

In summary, the travel time Estimation project represents a significant advance in automobile optimization; This progress is also highlighted by its commitment to accuracy, precision, F1 scores and return measurements. Through rigorous data analysis, modeling and evaluation, the project achieved a robust forecasting process that can provide timely and reliable forecasts.

The project focuses on accuracy to ensure commuters, truck drivers and urban planners can rely on predictive analytics to inform planning, scheduling and resource allocation decisions. The system provides better and more efficient transportation by reducing estimation errors and deviations from actual travel times, increasing user confidence and satisfaction.

Additionally, the project focuses on accuracy to ensure that downsides such as time estimation errors are minimized, thus reducing the potential impact and delays in transportation. This fact is important for important tasks in transportation, such as logistics planning and public transport management; Actually, time is very important.

The combination of F1 scores and retrospective testing further increases the reliability and validity of the program. By measuring accuracy and feedback, the system can detect positive (correctly estimated travel time) and negative (reported delays) and minimize negatives (inaccuracies, including estimates) and negatives (inaccurate estimates).

Overall, the Travel Time Prediction project marks a transformative leap forward in the field of traffic forecasting, providing unparalleled accuracy, precision and confidence in travel time prediction. By using state-of-the-art measurement methods and committing to continuous improvement, the project is paving the way for smarter, more efficient and easier transportation, ultimately improving mobility, connectivity and quality of life for people and communities around the world.

9. Reference

[1]Siripanpornchana, Chaiyaphum, Sooksan Panichpapiboon, and Pimwadee Chaovalit.

"Travel-time prediction with deep learning."

2016 ieee region 10 conference (tencon). IEEE, 2016.

[2] Siripanpornchana, C., Panichpapiboon, S., & Chaovalit, P. (2016, November). Travel-time prediction with deep learning. In 2016 ieee region 10 conference (tencon) (pp. 1859-1862). IEEE.

[3] Liu, Yangdong, et al. "Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models." 2017 IEEE 20th international conference on intelligent transportation systems (ITSC). IEEE, 2017.

[4] Liu, Y., Wang, Y., Yang, X., & Zhang, L. (2017, October). Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models. In 2017 IEEE 20th international conference on intelligent transportation systems (ITSC) (pp. 1-8). IEEE.

[5] Abdollahi, Mohammad, Tannaz Khaleghi, and Kai Yang. "An integrated feature learning approach using deep learning for travel time prediction." Expert Systems with Applications 139 (2020): 112864.

[6] Abdollahi, M., Khaleghi, T., & Yang, K. (2020). An integrated feature learning approach using deep learning for travel time prediction. Expert Systems with Applications, 139, 112864.

[7] Servos, Nikolaos, et al. "Travel time prediction in a multimodal freight transport relation using machine learning algorithms." Logistics 4.1 (2019): 1.

[8] Servos, N., Liu, X., Teucke, M., & Freitag, M. (2019). Travel time prediction in a multimodal freight transport relation using machine learning algorithms. Logistics, 4(1), 1.

[9] Qiu, Bo, and Wei Fan. "Machine learning based short-term travel time prediction: Numerical results and

comparative analyses." Sustainability 13.13 (2021): 7454.

[10] Goudarzi, Forough. "Travel time prediction: comparison of machine learning algorithms in a case study." 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE, 2018.

[11] Goudarzi, F. (2018, June). Travel time prediction: comparison of machine learning algorithms in a case study. In 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 1404-1407). IEEE.

[12] Deb, Bilash, et al. "Travel time prediction using machine learning and weather impact on traffic conditions." 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). IEEE, 2019.

[13] Deb, B., Khan, S. R., Hasan, K. T., Khan, A. H., & Alam, M. A. (2019, March). Travel time prediction using machine learning and weather impact on traffic conditions. In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT) (pp. 1-8). IEEE.

[14] Wang, Zheng, Kun Fu, and Jieping Ye. "Learning to estimate the travel time." Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2018.

[15] Ahmed, I., Kumara, I., Reshadat, V., Kayes, A. S. M., van den Heuvel, W. J., & Tamburri, D. A. (2021). Travel time prediction and explanation with spatio-temporal features: A comparative study. Electronics, 11(1), 106.